A Sensor Device for Automatic Food Lifelogging that is Embedded in Home Ceiling Light: A Preliminary Investigation

Takuya Maekawa

Grad. Sch. of Information Science and Technology, Osaka University

2-1 Yamadaoka, Suita, Osaka, 565-0871, JAPAN

maekawa@ist.osaka-u.ac.jp

Abstract—**Due to the recent proliferation of digital cameras and smart phones with camera devices, many researchers have attempted to store and analyze photographs (images) that capture a user's meal. Simply storing photographs of meals before eating can encourage weight loss. Also, by analyzing the images, some researchers attempt to estimate the nutritional composition of the meal. However, these approaches rely on images manually photographed by the user. So, when the user forgets to take a picture of his/her meal, the information related to the meal will be lost. In this paper, we propose and design a sensor device with a camera that automatically takes a photograph of a user's meal. The device is attached to a ceiling light in a dining room of the user's house. So, the device is supplied with electricity from the light. Also, the device has a camera and uses it to capture the dining table under the ceiling light. With this device, we can automatically and continually take photographs during the user's mealtime. Here the problem is how to determine a representative photograph of the user's meal from the continually captured images. In this paper, we investigate how to find the representative photograph captured during the mealtime by using the complexity of an image.**

I. INTRODUCTION

Since many health problems are caused by our daily behaviors (e.g., smoking, insufficient exercise, overeating, or unbalanced diet), many researchers in the ubiquitous and pervasive computing research fields have attempted to sense and record a user's daily activity information. In particular, dietary control by using photographs that capture a user's meals has received attention in the field of healthcare. The recent proliferation of digital cameras and smart phones with camera devices permits an end user to easily take and store photographs of his/her meal. With the images (photographs) of the user's meal, many dietary control studies have been conducted. Zepeda et. al. argue that photographic food diaries can alter attitudes and behaviors of a user associated with food choices [8]. Aizawa et. al. attempt to estimate the nutritional composition of the meal by analyzing the photographs of the meal [1].

However, many studies related to the food lifelogging rely on images manually photographed by an end user by using a digital camera or a smart phone with a camera device. When the user forgets to take a picture of his/her meal, the information related to the meal will be lost. So, in this paper, we attempt to automatically obtain a photograph of the user's meal by using a sensor device installed in the user's house. Our proposed sensor device is attached to a ceiling of a dining room of the user's house. And, the device captures the dining table under the ceiling light with its camera. Here we assume that the device captures the user's meal every day, so the device must be supplied with electricity. (The device exhausts its battery very quickly if the device runs with its internal battery because the device employs a camera sensor device.) So, in this paper, we focus on a ceiling light in the dining room of the user's house. We design the device that is supplied with electricity from a socket adapter inserted between a home light bulb (fluorescent) and its socket.

During the mealtime, the ceiling light is turned on. That is, the device is supplied with electricity from the light. While the light is on, the device continually captures photographs of the user's meal. Here the problem is how to determine a representative photograph of the user's meal from the continually captured images. Before a user eats meals, he/she prepares the meals in the user's kitchen and brings the meals to the dining table. In many cases, the user first brings dishes to the table one after another and then he/she starts eating after he/she finished bringing all the dishes to the table. So, the device should find the time when the user finished bringing all the dishes to the table. The photograph that is captured at the time is the representative image of the meal. This is because a photograph that is captured after/before the time may not include several dishes or meals. Also, because we assume a small and cheap device (such as Arduino), the algorithm should be computationally inexpensive.

In the rest of this paper, we first introduce the design of our prototype sensor device. Then we investigate how to find a representative image of a user's meal by analyzing images captured by the device.

II. RELATED WORK

As mentioned in the introduction section, many studies try to analyze images that capture a user's meal. Aizawa et. al. attempt to estimate the nutritional composition of the meal by analyzing the photographs of the meal [1]. Kitamura et. al. try to estimate food balance by analyzing food images [5]. Nohara et. al. attempt to employ food lifelogs to enhance a medical care consult [6]. These studies utilize photographs manually captured by a user. Lifelogging systems with wearable cameras also relate to our study. Many studies try to capture daily life events (including mealtime) with wearable cameras such

Fig. 1. Our prototype device that is attached to a ceiling light.

as Microsoft SenseCam [2], [4], [3]. On the other hand, our device focuses only on meal events and installed in an environment of interest (ceiling light).

III. PROPOSED DEVICE

A. Design

As mentioned in the introduction section, we focus on a ceiling light in a dining room and design a sensor device that is attached to the ceiling light. So, we design the device that is supplied with electricity from a socket adapter inserted between a home light bulb (fluorescent) and its socket. Fig. 1 shows our prototype sensor device attached to a ceiling light socket. Also, Fig. 2 shows example images captured by the device. So, the dining table was captured from the ceiling by using the device.

Our designed device has the following features.

1) The device is turned on only when the ceiling light in the dining room is turned on. While the light is on (i.e., mealtime), the device continually captures images.
2) The device analyzes the captured images during the mealtime and finds the best one (i.e., representative image). The representative image is used for food lifelog applications. Note that, because our device is a prototype (based on Android phone), the device simply stores images in its SD memory card. Also, we are now developing an Arduino-based device.

B. Finding the best photograph

With our sensor device, we can obtain a sequence of images captured during a user's mealtime. From the images, we should find a representative image. Here, when the user prepares meals, the user brings the meals (dishes) to the dining table one after another. After bringing all the meals, the user starts eating the meals. We consider that new dishes are continually added to the captured images before the user finishes bringing all the dishes to the table. So, the complexity of an image may increase before the user finishes bringing all the dishes. After that, the user starts eating the meals. That is, the foods on the dining table will be consumed. So, the complexity of an image may gradually decrease. Based on the assumption, in this paper, we compute the complexity of a captured image

Fig. 2. Example images captured by our device.

and find an image with the highest complexity value. When an image includes many foods, the image also includes many edges. So, we compute the complexity value based on detected edges. The complexity of an image p is defined as follows [7].

$$K(p) = \sum_{i,j \in \{x,y\}} k(D_{i,j}),$$

where $D_{i,j}$ is an image obtained by employing a Sobel filter (an edge detector) for the j coordinate after employing for the i coordinate. For example, when we obtain $D_{x,y}$, we employ a Sobel filter for the x coordinate and then employ a Sobel filter for the y coordinate. Also, $k()$ is defined as follows.

$$k(p) = \sum_n p_n,$$

where p_n shows the pixel value of the nth pixel.

Fig. 3 shows a sequence of images captured by our device and a transition of complexity values of the images. As shown in the results, the complexity value increases as many dishes are placed on the table. The image at the 9th frame was the best because it was captured just before experimental participants started eating. However, the complexity value of the 23th image was the highest. As shown in Fig. 3, the complexity value did not decrease quickly after all the dishes were placed on the table. This is because the foods on the table are not consumed quickly. Also, after experimental participants started eating, their bodies and clothes increased the complexity value. We consider that we can find the best time (photograph) by detecting time when the complexity value first reaches $x\%(x < 100)$ of the highest complexity value. In our investigation, when x is about 90%, it works fine.
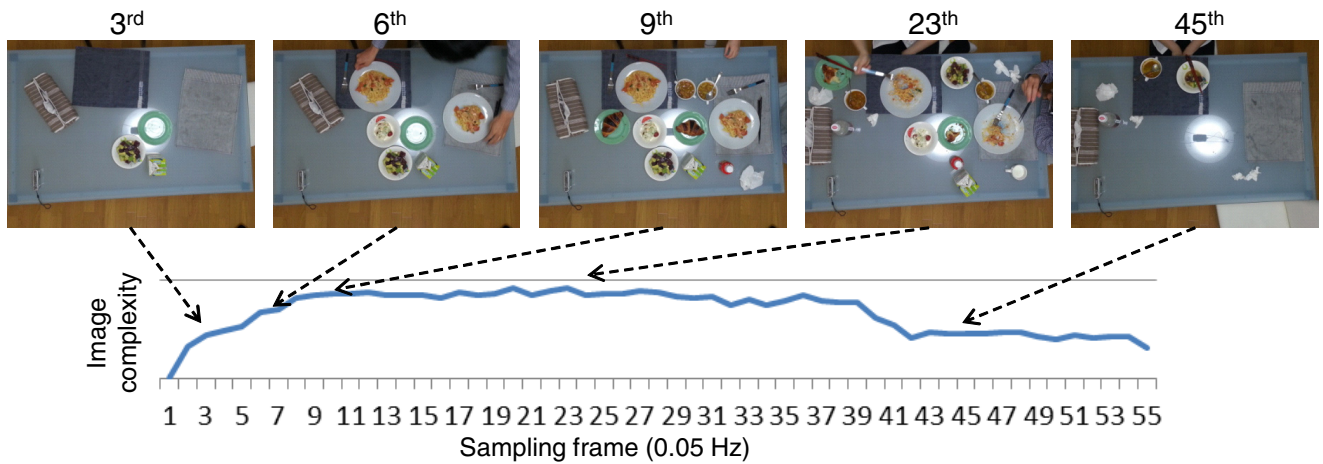
Fig. 3. Sequence of captured images and transition of complexity values.

Furthermore, we found that, before all the dishes were placed on the table, the experimental participants (residents) walked around the table to bring the dishes. After all the dishes were placed on the table, they did not walk around the table because they started eating. So, we consider that this phenomenon can be roughly detected by infrared sensors. By combining camera data and infrared sensor data, we may be able to find the best time with higher accuracy.

IV. CONCLUSION

In this paper, we propose a new sensor device equipped with a camera that automatically captures a user's meal. We design and introduce our prototype device and investigate how to find a representative image of a user's meal from a sequence of images that are captured during his/her mealtime. As a part of our future work, we plan to develop a sensor device that can find the representative image, and sends the image to a food lifelog application. (In this implementation, we analyze captured images off-line.) Also, we plan to employ other sensors such as infrared sensors to accurately find the representative image.

REFERENCES

[1] Kiyoharu Aizawa, Gamhewage C de Silva, Makoto Ogawa, and Yohei Sato. Food log by snapping and processing images. In *International Conference on Virtual Systems and Multimedia (VSMM 2010)*, pages 71–74, 2010.

[2] M. Blum, A.S. Pentland, and G. Tröster. Insense: Interest-based life logging. *IEEE Multimedia*, 13(4):40–48, 2006.

[3] Aiden R Doherty and Alan F Smeaton. Automatically segmenting lifelog data into events. In *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2008)*, pages 20–23, 2008.

[4] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood. Sensecam: A retrospective memory aid. In *Ubicomp 2006*, pages 177–193, 2006.

[5] Keigo Kitamura, Chaminda de Silva, Toshihiko Yamasaki, and Kiyoharu Aizawa. Image processing based approach to food balance analysis for personal food logging. In *IEEE International Conference on Multimedia and Expo (ICME 2010)*, pages 625–630, 2010.

[6] Mitsuo Nohara, Nobuhide Kotsuka, Masayuki Hashimoto, and Hiroki Horiuchi. A study on food-log application to a medical-care consult via telecommunications. In *International Conference on Virtual Systems and Multimedia (VSMM 2010)*, pages 88–91, 2010.

[7] Robert Rich. Image contrast, complexity, and stability. *Computer vision, graphics, and image processing*, 26(3):394–399, 1984.

[8] Lydia Zepeda and David Deal. Think before you eat: photographic food diaries as intervention tools to change dietary decision making and attitudes. *International Journal of Consumer Studies*, 32(6):692–698, 2008.